

REMARKS

Claims 1-12 and 24-30 remain in the application with claims 1, 5 and 24 having been amended hereby.

In the Abstract of the Disclosure, the examiner believes that the covered material is not supported by the specification.

The examiner correctly notes that the Abstract indicates that the invention is capable of separating music into voice and accompaniment, translating the voice to produce a translated voice and play the accompaniment along with the translated voice.

It is respectfully submitted that these functions are fully supported by the disclosure. Beginning on page 2, line 15 of the specification, a separating unit for separating the lyric (voice) information and the accompaniment information is described. This description is followed by a description of a processing unit capable of speech-recognition. This description is followed by a description of a processing unit capable of translating the recognized speech. This description is followed by a description of a speech synthesizing unit for synthesizing the translated speech information into sound that is presented with the accompaniment information to generate the final output of synthesis information.

It appears that the objection is based on the fact that the precise manner in which speech is separated, recognized, translated and synthesized is not specifically disclosed in the specification.

It is important to understand that by themselves, these four technological accomplishments are well understood in the art. More specifically, the technology for separating speech from musical accompaniment is well understood in the art. Speech recognition is well understood in the art. Text translation is well understood in the art. Voice synthesis is well understood in the art. The precise manner in which these technologies operate is not specifically disclosed in the specification but nor are they independently claimed. What is claimed is the Information Processing Method and Apparatus that utilizes all four of the above technologies in a novel way. It is this Information Processing Method and Apparatus that is fully disclosed in the specification. The abstract, as it currently stands, summarizes how the four above mentioned technologies, which are well known in the art, come together to form the current invention.

In light of the above reasoning, and the Amendments made, it is respectfully requested that the examiner reconsider his objection to the Abstract.

The examiner objects to the drawings under 37 CFR 1.83(a) as not showing every feature of the invention specified in the claims.

Figs. 4, 5 and 6 illustrates the process and apparatus for how music (D1) (musical number information) is broken into vocal information (D3), by the vocal extraction unit (212b), how the music is transformed into accompaniment information (D2) by the vocal canceling unit (212a), how the lyric information is extracted from the music by the recognition processing unit (321b), how the lyric information is translated into second language lyric information by the translation processing unit (321d) and how the second language lyric information is used to generate a speech signal data by the speech generating unit (322d).

The drawings therefore show every feature of the invention specified in the claims. It is therefore respectfully requested that the examiner reconsider the objection to the drawings as failing to satisfy 37 CFR 1.83(a).

The examiner contends that the changes to Figs. 4-6 only show a desired result without providing any useful showing of how the various steps of the present invention are accomplished.

This objection appears to be related to the examiner's objection to of the Abstract discussed above. It is respectfully restated that the four above-mentioned technologies; speech separation, speech recognition, text translation and speech synthesis, are well known to those with skill in the art. Disclosure of these technologies need not be individually enabling because these technologies are not individually claimed. It is the Information Processing Method and Apparatus that is specifically claimed, and it is this invention that is specifically enabled.

With respect to the previous Amendment, the examiner has objected to this amendment under 35 U.S.C. 132 for introducing new matter to the disclosure.

The Examiner specifically objected to "some sort of ID ... based on a subscription list and ... dependant on payment of a use fee" added to page 14 of the Specification.

It is respectfully submitted that the disclosure concerning the payment of a use fee begins on page 7, line 5 of the original disclosure, "This assessment is done via an assessment communication network 5 so that the fee is collected from the user. This assessment communication network 5 is constituted by, for example, the communication medium, such as a telephone network, with the server device

1 being connected via the assessment communication network
5 to a computer device of banking facilities which have
made contract in connection with payment of the use fee of
the information distribution system."

The Examiner contends that "no such definition of
collation data and permission to use a terminal device was
previously given." It is respectfully submitted that the
disclosure concerning the ID, collation data and permission
to use the terminal device begins on the page 14, line 10
of the original disclosure:

A collation processing unit 104 collates the
terminal ID data of the portable terminal device
3 sent along with the request information to the
terminal ID data of the portable terminal device
that is currently able to use the information
distribution system (stored as user-related data
in the storage unit 102) to send the results of
collation to the controller 101. Based on the
results of collation, the controller then decides
whether the information distribution system is or
is not permitted to be used to the portable
terminal device 3 loaded on the intermediate
transmission device 2 of the destination of
transmission of the request information.

In light of this reasoning, it is respectfully
requested that the examiner reconsider his objection to the
previous Amendment as containing new matter as defined in
35 U.S.C. 132.

In the claims, the Examiner has rejected all claims
remaining in the Application (1-12 and 24-30) under 35

U.S.C § 112, second paragraph, as being indefinite for failing to point out and distinctly claim the subject matter which applicant regards as the invention.

With respect to claims 1 and 24, the Examiner contends that the new claim language, specifically, "first vocal-containing musical number information" contradicts language in the specification, "Generating the first language lyric information by speech recognition of the first vocal information." The "first vocal-containing musical number information" is here referred to in the specification, in its abbreviated form, as the "first vocal information."

The examiner contends that "the lyrics must be derived from speech" and yet the musical information is not limited to music containing speech. While it is true that not all forms of music contain lyrics, non-lyrical music is still "musical information," however, the fact that there is little or no lyrics to process does not change how the music is treated according to the claimed invention, and this is perfectly in-line with the current disclosure.

The disclosure on the original page 23 does talk about "karaoke information" and the claims talks about the same information in terms of "accompaniment information" however, the use of these distinct terms is not confusing as both terms are discussed as "D2" in the specification.

The use of the term "karaoke information" is used to add depth to the disclosure.

The specification does not say that the karaoke information is a substitute for traditional forms of stereo encoding as the Examiner suggests, the Specification only illustrates that the karaoke information is obtained from canceling the speech signal (vocal information) and that some well known techniques for speech canceling relate to finding signals that are "fixed at the center on stereo reproduction, that is, sounds that are as loud in the right channel as they are in the left channel. This disclosure merely points to a method of speech canceling well known in the art and should not be read as somehow limiting the disclosed invention.

The Examiner points out many types of music are ill suited for use with the current invention. The Examiner specifically refers to choral music with 4-part harmony where multiple singers simultaneously sing different lyrics. The Examiner's point is well taken, however, the Examiner merely points out that certain forms of music are ill-suited for karaoke use. Additionally, nothing in the disclosure or the claims limits the separation, canceling and recognition of lyrics to music where there is a lone voice. The arts of speech recognition, speech separation

and speech canceling are not limited to the recognition, separation and canceling of one voice at a time and the current invention is also not so limited. It is respectfully pointed out that the intricacies of speech recognition, speech separation and speech canceling are not the subject of the current invention.

The Examiner contends that the phrase "language lyric information" is not clearly linked to either part of the "vocal information D3" or part of the "karaoke information D2." Fig. 5, as amended, clearly establishes that vocal information is sent to the Sound Analysis Unit 321a, then to the Recognition Processing Unit 321b, then to the Translation Processing Unit 321d, finally to emerge as Language Lyric Information. The process by which this occurs is clearly described in the specification and recited in the claims.

The Examiner contends that while the only specific application mentioned in the specification is for karaoke related data, the claims are broad enough to include a wider range of multimedia information. Claim 1 refers to the input as "the first vocal-containing musical number information" therefore limiting the scope of the claim to music containing vocals. While the input data must have music containing vocals, it is certainly possible that such

data is accompanied by, for example, video or picture data, as described in the first page of the Specification. It is quite common in the art to have karaoke accompanied with video and/or picture data, and such data is covered in the disclosure and the claims are worded in such a way as to include this possibility.

With respect to the Examiner's rejection of claims 1-12 and 24-30 under 35 U.S.C. 112, first paragraph as not being enabling, the Examiner contends that the specification allows for the input of diverse forms of media such as "audio, text and picture information." The Examiner contends that while the data that can be input is broad, the disclosure only teaches how a particular form of data is handled, that is, music containing vocal information.

The Examiner contends that someone with ordinary skill in the art could not hope to predict how diverse forms of media will be processed by the current invention. The Examiner fails to consider, however, that among the relevant arts to consider is the art of karaoke. One skilled in the art of karaoke would know exactly how picture data such as picture data and text data is treated by the current invention. It is well known in the art of karaoke that often while karaoke music plays, text and pictures are displayed making for a more entertaining

karaoke experience.

The Examiner objects that there is no particular limitation on the form of the communication network used to transfer data within the distribution system and therefore this disclosure "does nothing to define the data or its components."

The text of 35 U.S.C. 112, first paragraph, that the Examiner has quoted holds that the "specification shall contain a written description of the invention, and of the manner and process of making and using it..." The method of data transmission is not relevant to the proper function of the current invention, and therefore the method of data transmission to be used has not been limited. Several example methods of data transmission have been provided to suggest some particular ways in which data can be transmitted, however, the method of transmission is not an essential feature of the invention and therefore any method of data transmission can be used. The description of the method of data transition is therefore not meant to define the data or its components. It is meant to illustrate that the invention can be constructed using any available method of data transmission.

With respect to the intermediate transmission device, all of its essential features have been clearly described

in the specification and recited in the claims. What is significant is that the terminal device "mainly has the function of receiving the information sent mainly from the server device 1 by a communication control terminal 201 and outputting the received information to the portable terminal device 3. The intermediate transmission device 2 also has a charging circuit for electrically charging the portable terminal device 2" (beginning on page 7, line 15).

The construction of the intermediate terminal device is well described in that, by reference to its key features, a person skilled in the art would be enabled to construct the intermediate terminal device. Because this device, by itself, is not the claimed invention, its construction is already known in the pertinent art.

Additionally there are many such components of the current invention that are well known in the art. Among them is how to construct a computer network for the communication of data. The reference to the TCP/IP protocol illustrates one protocol for communicating data, known in the art. The disclosure clearly states that any such protocol may be used. The examiner contends that because the form of the data to be transmitted is unknown, it is not clear what data transition protocols are acceptable. The Examiner is respectfully reminded that data transfer

protocols such as TCP/IP are useful for communicating all digital communications regardless of whether the digital data is representative of pictures, sounds or anything else.

The same argument is equally applicable to the various methods of data compression that may be used to facilitate the transfer of digital data. Twin VQ is merely one available method for compressing data. It is respectfully submitted that the method of data compression and communication have only an ancillary roll in the present invention, and the methods selected have no bearing on the applicability of the current invention.

The Examiner contends that the "original Specification, page 14, second full paragraph failed to indicate what data is "collated." The examiner believes that it is not apparent what data is collated and how the collated data is used.

While this paragraph from the original specification is concededly oddly phrased, the first Amendment to the specification clarifies the intended meaning of this paragraph. The intended meaning of this paragraph has been further clarified in the most recent Amendment. According to these clarifications, the "collation processing unit 104 collates the terminal ID data of the portable terminal

device 3 with the terminal ID data of the devices currently able to use the information distribution system." Although its original form was concededly oddly worded, only the form of this paragraph has changed. Its substance remains the same. The collation processing unit makes sure the terminal ID of the device requesting data is authorized to receive the data requested.

The Examiner contends that the description of the payment of a use fee is new matter. The Examiner believed that the asserted foundation for this use fee is established by the mention of "a charging circuit, for supplying the power to the various parts." Concededly, this charging circuit relates to electrical charging and not billing. It is respectfully submitted, however, that the foundation for the payment of a use fee is disclosed elsewhere in the specification as originally filed. The Examiner is requested to look at the previously quoted portion of the original specification beginning on page 7, line 5. "This assessment is done via an assessment communication network 5 so that the fee is collected from the user. This assessment communication network 5 is constituted by, for example, the communication medium, such as a telephone network, with the server device 1 being connected via the assessment communication network 5 to a

computer device of banking facilities which have made contract in connection with payment of the use fee of the information distribution system." It is respectfully requested that the Examiner reconsider his classification of the previous Amendment as being new matter with respect to the assessment of fees.

The Examiner contends that the last paragraph of page 14 (continuing to page 15), as originally written, is confusing. While it is conceded that the original language was difficult to understand, the same paragraph, as amended by the previous Amendment, has corrected the confusing language. It is respectfully requested that the Examiner take another look at the paragraph as Amended in the prior Amendment, as it is believed that this Amendment is responsive to the Examiner's concerns.

The Examiner believes that the current specification provides no evidence that the invention can perform vocal separation, speech recognition, text translation and speech syntheses, as described in the specification and claimed in the claims. The Examiner is apparently concerned that few details are provided for carrying out these aspects of the present invention.

As discussed above, these four technological accomplishments are well understood in the art. More

specifically, the technology for separating speech from musical accompaniment is well understood in the art. Speech recognition is well understood in the art. Text translation is well understood in the art. Voice synthesis is well understood in the art. The precise manner in which these technologies operate is not specifically disclosed in the specification but nor are they independently claimed. What is claimed is the Information Processing Method and Apparatus that utilizes all four of the above technologies in a novel way. It is this Information Processing Method and Apparatus that is fully disclosed in the specification.

The Examiner reiterates his concern that the Specification holds that the vocal canceling unit can use can use a well-known technique of canceling the speech signals fixed at the center on stereo reproduction, while the specification does not limit the input music to a type utilizing stereo separation. This well-known technique was offered as an example of one technique for canceling vocals, and according to the specification, other well-known techniques may be used. Additionally, the Examiner is apparently making the assumption that the discussed well-known technique for canceling speech signals fixed at the center on stereo reproduction would be ineffective when presented with monophonically recorded music. Methods for

vocal cancellation are well known in the art, and have overcome the problems the examiner has mentioned, the specifics of which are not expounded upon in the current specification as it is presumed that the application of this invention relies on the knowledge of those skilled in the relevant arts.

This same explanation is well suited to the Examiner's contention that the current specification does not adequately disclose how speech synthesis is achieved. The Examiner correctly points out that changes in pitch can change the meaning of words in certain languages (the Examiner notes Chinese as an example) and that these nuances pose a particular problem for voice recognition that the instant specification does not address. Again, the use of speech synthesis is well known in the art and these methods are not limited to monotonic languages.

In light of the above remarks, and the amendments made, it is respectfully submitted that a INFORMATION PROCESSING APPARATUS AND METHOD, as taught by the present invention and as recited in the amended claims, is adequately described and enabling in satisfaction of 35 U.S.C. 112, first paragraph.

With respect to the Examiner's rejection of claims 1-12 and 24-30 under 35 U.S.C. § 103, first paragraph as being

unpatentable over Stelovsky (5,613,909) in view of Bordeaux (4,852,170) and Lyberg (5,546,500), the current invention claims a method and system for inputting music containing voice, separating the music into voice and musical accompaniment, employing speech recognition processing on the voice to generate lyrical information, translating the lyrical information into a language other than the language that the music was originally sung in and using speech synthesis to create vocal information in the second language.

Stelovsky teaches traditional karaoke wherein each element of the music is recorded onto media as a separate track. For example, there is a music video track, an accompaniment track, a vocal track and a text track.

Stelovsky also suggests that "such tracks can be generated directly using a software program rather than using digitized or analog data. The audio track can be generated rather than recorded (e.g. using a speech generator.) The textual track can be generated rather than recorded (e.g. using a direct translation into another language.)"

Stelovsky does not show or suggest that a separation unit can be used to separate voice from accompaniment or that speech recognition can be used to generate the text

track.

The Examiner believes that Stelovsky teaches the separation of lyrics and accompanying music. While Stelovsky does, however, advocate the separate use of lyrics and accompanying music, the recorded media used in Stelovsky contain separate tracks for lyrics and accompanying music. Therefore, lyrics and music come pre-separated and are prerecorded in discrete tracks.

Bordeaux relates to speech recognition.

Lyberg relates to a device for receiving speech in a first language, translating the received speech into a second language and generating speech in the second language.

None of the above-cited references teach separating voice from accompanying music. Stelovsky teaches using a prerecorded medium containing separate tracks for voice and accompanying music. Stelovsky also suggests that "such tracks can be generated rather than being recorded," however, in light of the given example "The textual track can be generated rather than recorded (e.g. using a direct translation into another language.)" Stelovsky is not suggesting that a voice track and an accompaniment track can be generated from a single recorded track, Stelovsky is suggesting that a first text track in one language can be

translated into a second text track in another language without that second text track ever having to be recorded. Stelovsky is suggesting that certain audio data need not be recorded at all but may be generated. In the current invention, audio data is broken into voice and accompaniment components. This breaking down of a prerecorded audio is neither suggested nor taught in Stelovsky. The track generation spoken about in Stelovsky more closely parallels the second language lyric synthesis performed in the current invention. While it is conceded that Stelovsky teaches voice synthesis, Stelovsky does not teach or suggest that a separation unit can be used to separate voice from accompaniment.

Stelovsky also does not teach or suggest that speech recognition can be used to generate text. Bordeaux and Lyberg each teach the use of speech recognition. The mere fact, however, that the prior art may be modified in the manner suggested by the Examiner does not make the modification obvious unless the prior art itself suggests the desirability of the modification. See In re Fritch, 972 F.2d 1260, 1266, 23 U.S.P.Q.2d 1780, 1783 (Fed. Cir. 1992). The motivation to modify cannot come from the present invention. See Heidelberger Drucksmaschinen AG v. Hantscho Commercial Products, 21 F.3d 1068, 1072, 30 U.S.P.Q.2d


1377, 1380 (Fed. Cir. 1994). Because Stelovsky does not suggest the use of speech recognition, speech recognition taught in Bordeaux and Lyberg may not be so combined.

In light of the above remarks, and the amendments made, it is respectfully submitted that a INFORMATION PROCESSING APPARATUS AND METHOD, as taught by the present invention and as recited in the amended claims, is not unpatentable under 35 U.S.C. § 103, first paragraph over Stelovsky (5,613,909) in view of Bordeaux (4,852,170) and Lyberg (5,546,500).

Favorable reconsideration is earnestly solicited.

Respectfully submitted,

COOPER & DUNHAM LLP


Jay H. Maioli
Reg. No. 27, 213

JHM/JBG